DOCUMENT RESUME

ED 279 398 PS 016 308

AUTHOR Stallings, Jane; Krasavage, Eileen
TITLE Peaks, Valleys, and Plateaus in Program Implementation: A Longitudinal Study of a Madeline Hurter Follow Through Project (1985 Data).
PUB DATE Jan 86
NOTE 75p.; Paper presented at the Family and Research Action V Conference (Lubbock, TX, January 22-24, 1986).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS *Academic Achievement; Classroom Research; Classroom Techniques; Compensatory Education; Cost Effectiveness; *Disadvantaged Youth; Elementary Education; *Elementary School Students; *Instructional Development; Kindergarten; Longitudinal Studies; Mathematics; Models; Program Effectiveness; Program Implementation; Reading; *Teaching Methods; *Time Factors (Learning)
IDENTIFIERS Instructional Theory into Practice; Instructional Variables; *Project Follow Through

ABSTRACT

This paper reports findings from the National Institute of Education's Napa County Project, one of four follow through projects funded over a period of four years to develop and implement instructional programs that would improve teaching practices, increase students' academic engaged rates, and increase gain on reading and mathematics achievement tests. The assumptions of the follow through project were that effective staff development would lead to improved instructional practice and the improved instructional practice would help students be more productively engaged in academic tasks. In the Napa County Project, which involved two schools with the largest percentage of children eligible for Chapter I in their districts, Madeline Hunter's Instructional Theory into Practice (ITIP) was selected as the training model. After a planning phase, begun in 1981, the project implemented the training program, collected implementation and student data, and revised the training program on the basis of yearly evaluations. Instruments to measure program implementation and rate of students engaged were developed and piloted in kindergarten through fourth grade classrooms. Two additional schools were selected to provide a basis for comparisons. Results of numerous analyses indicated that the primary impact of the program culminated at the end of the second year of implementation. Reasons for the subsequent decline are briefly discussed. (RH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED279398

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Peaks, Valleys, and Plateaus in Program Implementation: A Longitudinal Study of a Madeline Hunter Follow Through Project (1985 Data)

Jane Stallings
Eileen Krasavage
1986(in press)

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Jane Stallings

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

University of Houston
University Park

PS 016308

BEST COPY AVAILABLE

Peaks, Valleys, and Plateaus in Program Implementation:
A Longitudinal Study of a Madeline Hunter Follow Through Project

The project described here was one of four Follow Through Projects funded by the National Institute of Education (NIE) in 1981. NIE's purpose for the projects was to develop and implement instructional programs which would improve the practice of teaching, increase students' academic engaged rates, and increase gain on reading and mathematics achievement tests. The projects funded were located in Detroit, Michigan; Oakland, California; Cotapaxi, Colorado; and Napa County, California. This paper reports findings from the Napa County project. Each project was required to select two project schools in the districts which served the largest number of Title I children. The program was to be delivered to all kindergarten through fourth grade classrooms in each of the schools.. To examine the premise that improved classroom practices alone would positively effect student engaged rate and achievement, other sources of federal or state aid and assistance were withheld from the classrooms in the project. Title I classroom aides were reassigned to upper grade classrooms or other schools. A central question to be addressed by the study was the cost effectiveness of comprehensive and continuous staff development programs compared to the cost and effect of classroom aides and other Title I resources.

The projects were funded over a period of four years. The planning phase started in the spring of 1981 when schools were recruited. Madeline Hunter's Instructional Theory into Practice (ITIP) was selected as the training model for the Napa County

project. (See Hunter, 1976 for a description of the model). Instruments to measure ITIP program implementation and student engaged rate were developed and piloted in kindergarten through fourth grade classrooms. The purposes of Phases II, III, and IV were to implement the training program, collect implementation and student data, and revise the training program on the basis of the yearly evaluations. The findings from the overall study are reported in this article.

METHOD

The project was conducted in two central California school districts. It involved the principals, first through fourth grade teachers, and students at two elementary schools. The Follow Through project staff included a director, a coordinator, and a trainer who were all qualified trainers in the Madeline Hunter ITIP model. Although classroom management and effective use of time were included as training components, ITIP was the major training component during Phase II reported in Stallings, 1985 and Phase III reported in Stallings et. al, 1986. Relationships between program implementation, teacher behavior, and student outcomes were examined over three phases.

Sample

The two schools selected for the study had the largest percentage of Chapter I eligible children in their school districts. Locating control schools was difficult since the two school districts are small and the two project schools were selected because they served children scoring the lowest on tests and coming from the lowest socioeconomic homes. The two control schools selected were the best available according to these two

criteria. The total number of students in Treatment School #1 and its control were very similar. There were approximately 450 students each year of the project. Treatment School #2 was less than half the size of its control school.

To compare the background features of the treatment and control schools' populations, the student data collected in the California Assessment Program (CAP) have been used. These data were collected from third and sixth grade children each year. Since ours was a longitudinal study, these data allowed us to examine trends as the children from first grade through fourth grade passed through the CAP testing program..

The background factors of each schools' population determined a level of expected school performance. The background factors included socioeconomic status (SES). For grades three and six, parents' occupation was used as an indicator of SES. The parents' occupations were assigned values of 1, 2, or 3 (from unskilled to professional), and averages for a district or a school were computed. The percent of children from homes receiving Aid to Families with Dependent Children (AFDC) was also used in the calculation as was the percentage of children who were designated as limited English speaking (LES). Table 1 allows the comparison of treatment and control schools on these factors at each school site.

In School #1 the SES level for third graders ranged from the 13th percentile to the 30th percentile rank for California schools. The control school's lowest SES level was the 30th percentile and ranged as high as the 59th percentile in 1984. Over the four years, the treatment school served more AFDC and LES children in

Table 1

Background Factors of Project Schools

Schools	Year	N	SES*	Rank*	%AFDC*	Rank*	%LES*	Rank*
#1-Treatment								
Grade 3	1981-1982	52	1.83	29	28.2	87	5.1	51
	1982-1983	56	1.76	26	32.4	90	17.9	97
	1983-1984	69	1.52	13	32.4	90	7.2	59
	1984-1985	69	1.77	30	33.3	90	4.3	47
Grade 6	1981-1982	57	1.91	34	28.2	87	1.7	46
	1982-1983	61	1.95	40	32.4	90	4.9	65
	1983-1984	66	1.54	12	32.4	90	3.0	52
	1984-1985	44	2.47	84	33.3	90	0.0	1
#1-Control								
Grade 3	1981-1982	56	1.96	39	16.3	66	7.0	61
	1982-1983	54	1.81	30	14.4	59	0.0	1
	1983-1984	55	2.16	59	14.4	59	3.6	46
	1984-1985	54	1.81	33	21.2	74	0.0	1
Grade 6	1981-1982	63	2.05	46	16.3	67	4.8	66
	1982-1983	61	2.11	55	14.4	61	1.6	44
	1983-1984	66	2.08	54	14.4	60	0.0	1
	1984-1985	55	1.80	30	21.2	75	0.0	1
#2-Treatment								
Grade 3	1981-1982	25	1.54	12	29.5	88	18.5	82
	1982-1983	31	1.76	26	28.3	86	17.1	78
	1983-1984	32	1.52	13	28.3	86	3.0	43
	1984-1985	23	1.28	4	17.0	64	8.0	60
Grade 6	1981-1982	39	1.87	31	29.5	88	5.1	67
	1982-1983	30	2.00	44	28.3	87	12.5	84
	1983-1984	25	1.63	17	28.3	87	11.1	80
	1984-1985	23	1.57	15	17.0	65	8.7	72
#2-Control								
Grade 3	1981-1982	71	1.97	39	17.0	68	4.2	51
	1982-1983	68	2.03	48	21.4	77	1.5	33
	1983-1984	64	1.95	41	21.4	77	6.3	56
	1984-1985	75	1.99	47	20.9	73	3.9	46
Grade 6	1981-1982	111	2.15	55	17.0	69	3.6	61
	1982-1983	96	2.22	64	21.4	78	1.0	37
	1983-1984	84	2.22	66	21.4	77	4.7	61
	1984-1985	79	2.01	50	20.9	74	2.5	47

*Legend: SES = Socioeconomic Status Ratings
Rank = Percentile rank in State of California schools
AFDC = Aid to Families with Dependent Children (Percent Receiving)
LES = Limited English Speaking (Percent)

both third and sixth grades than did the control school.

Treatment School #2 tested about one-third as many students in first and third grade compared to the control school. There was, in fact, only one third grade classroom in the treatment school. The SES level of Treatment School #2 was lower at each year of comparison in both grade levels. The SES level of School #2 was lower in 1985 than in other years of the study. This population shift was accompanied by an increase in the percent of LES children served in the third grade. Approximately one-third of the children in this school were new each year.

Allocated Time

The time allocated to reading and mathematics differed somewhat for the two treatment schools. School #2 prescribed 70 minutes of reading and 50 minutes of mathematics for first through fourth grade. In School #1, time for first, second, and third grade was the same as in School #2. In fourth grade reading was reduced by 20 minutes and mathematics was increased by 10 minutes.

Staff Development

The staff development program focused on improving the instruction and classroom management skills of teachers. The expectation was that students would then be more engaged in their studies and improve their achievement test scores. To reach the goals of the project, a three-year staff development program for teachers and principals was developed.

Starting in January, 1983 during Phase II, there were five one-day training sessions to teach ITIP. These addressed the specific skills learned up to and including the five-step lesson plan (set, instruction, guided practice, closure, and independent

practice). These sessions were followed by classroom visits and feedback from the Follow Through project staff. During this period all teachers received at least five follow-up visits. Teachers who requested more assistance received as many as nine follow-up visits.

In Phase III, training activities were provided from August through April. These activities included: classroom management, effective directions; grouping; monitoring (praise, prompt, and leave); test-taking strategies; motivation, retention, practice theory--(ITIP); five-step lesson review and practice--(ITIP).

At the end of each inservice session, participants could schedule follow-up visits with the consultants conducting the workshops. These visits enabled participants to get feedback on their use of newly-learned strategies in their classrooms. The overall goal was to assist teachers in using the strategies which were expected to improve student engaged time and achievement.

The purpose of Phase IV was to encourage the principals and faculties at each treatment school to carry out the Follow Through program with minimal leadership from the project staff. It was meant to be a transition year. The focus of the evaluation was to see how the program was maintained in terms of teacher and student behavior. While there were only three days of release time for project inservice, the Follow Through staff was on-site weekly, responding to teachers' requests for assistance. Essentially, the fourth phase was a maintenance year.

Training and Support Services for Principals

The principals' training program was delivered through consultation, modeling, coaching, and exposure to a collegial

support group. Some of the training topics for principals included: characteristics of effective schools, leadership styles, clinical supervision, cognitive coaching, team building, collaborative planning, and effective meeting management.

Measurement

The instruments used to measure the quality and quantity of program implementation and student engaged rates were developed, piloted, and revised in Phase I of the project. These included the Instructional Skills Observation Instrument (ISOI), the Time Off Task observation (TOT), teacher and principal questionnaires, and interviews. Reading and mathematics achievement tests were available from the schools for all grade levels each year of the study.

Instructional skills observation. Teachers' implementation of the training program was measured by the ISOI (Wolfe, 1982). This instrument measured the quality and appropriateness of four program components: set/purpose/objectives; instruction; guided practice; and independent practice. The fifth step, closure, was subsumed in guided practice. Operational definitions for the observed components were developed with the assistance of Madeline Hunter. The critical issue was not whether or not a teacher used an element, but rather whether or not the element was needed for the lesson and if so, if it was used effectively. If a teacher accomplished each component perfectly, a score of 88 could be earned. Observers trained in the Hunter Program made the judgment of appropriateness and effectiveness. These observers were trained in a three-day session, where they achieved an interrater agreement of .83 or greater.

Each teacher was observed in two mathematics periods and two reading periods before the training was provided (to provide baseline data) and at each spring data collection period. Scores on the four subcategories were aggregated to a total score for each teacher. The total scores, averaged for the two reading observations, were used in the analyses. The same process was used for mathematics.

Engaged rate observations. Student engaged rate was measured for every student in the study using the Time Off Task Observation System. The sample of behavior included approximately 20 observations during reading and 20 during mathematics collected on two days at each data collection period (spring 1983, spring 1984, and spring 1985). A record was kept of each student who exhibited the following off-task behaviors during the observation period: chatting, disrupting, waiting, personal needs, and uninvolved. If a student was on task and doing what was expected, no entry was made beside the student's name. Computations for percentage student engaged rates began with the off-task frequency divided by the number of observations. To maintain unidimensionality of scaling, each student's percentages of time off-task for reading and for mathematics were subtracted from 100 to obtain an engaged rate. These percentages were then divided by 100 to allow for arcsin conversion.

Observers were trained during a two-day session that included an examination of interrater agreement in classrooms. The agreement among the three observers collecting data was .85 or above.

Achievement tests. Achievement tests were administered to

the students by each teacher in May each year of the study. School #1's district used the Stanford Achievement Test, and School #2's district used the California Achievement Test. It was not economically feasible for either school to change tests; thus, a percentile score was recorded for each student. This provided a standardized score that could be used to compare scores of students from the two schools. Percentile scores were transformed to normal curve equivalents for use in analyses.

Interviews. In spring of 1983, 1984, and 1985 the principal investigator had a structured interview with each teacher. The questions focused upon the use of the Five Step Lesson Plan.

ANALYSIS AND RESULTS

The assumption of the Follow Through project was that effective staff development would lead to improved instructional practice and the improved instructional practice would help students be more productively involved in academic tasks. Increased student involvement was expected to increase student achievement on standardized tests of reading and mathematics. The following analyses of impact on teacher behavior were conducted at the project level using student mean scores and at the teacher level using class means.

The Impact of Training Upon Teacher Behavior

Because the assumption was that the change process starts with the teachers, we have elected to present the teacher behavior change data first. These data report how the teachers performed on four components of the ISOI (i.e., set, instruction, guided practice, and independent practice). An intercorrelation of these components indicated a significant ($p<.05$) correlation among the

components in both reading and mathematics. The only exception (non-significant correlation) was between instruction and independent practice in mathematics.

Using the total ISOI scores for all teachers, one-tailed t-tests were computed to test the difference in teacher behavior from baseline through each phase of the study. Table 2 indicates there were significant positive changes in teacher behavior from baseline to spring 1983 and from spring 1983 to 1984 in reading and mathematics. The surprise is the reversal of this pattern in spring 1985. As a group, the scores of the 11 teachers remaining in the study dropped 11-12 points from 1984 to 1985 in both subject areas. Figures 1 and 2 graphically portray the fluctuations of these teacher scores. The solid lines identify the scores for the teachers in School #1 and the broken lines identify the teachers in School #2. The two teachers (H and M) who left the project in 1984 were in the lower half of performance on the ISOI for that year; thus, their absence from the 1985 data does not help explain the lower scores. Teacher G's scores are very erratic in mathematics and in 1985 certainly pulled the group mean down. However, seven of the 11 teachers' reading and mathematics ISOI scores decreased in 1985.

A possible explanation for the drop in scores is the lateness of collecting the observation data. Although we had planned to collect the data in early April to mid May, changes in school testing schedules extended our observation schedule until the end of May and the first of June. By this time in the year teachers tend to provide less and less formal instruction.

Table 2

T-tests on Mean Teacher ISOI Skills

Variables	Mean	Standard Deviation	Difference of Means	T-value	Degree of Freedom	One-tailed Probability
READING						
ISOI Winter 1983	45	10.2	10	1.93	12	*.038
ISOI Spring 1983	55	6.8				
ISOI Spring 1983	55	6.8	19	4.74	12	*.0005
ISOI Spring 1984	74	4.3				
ISOI Spring 1984	74	9.9	-12	-2.50	10	*.016
ISOI Spring 1985	62	21.1				
MATH						
ISOI Winter 1983	44	6.9	16	2.20	12	*.025
ISOI Spring 1983	60	8.4				
ISOI Spring 1983	60	8.4	16	2.88	12	*.008
ISOI Spring 1984	76	4.0				
ISOI Spring 1984	76	8.1	-11	-2.49	10	*.016
ISOI Spring 1985	65	16.3				

ISOI means were computed using raw scores.

TEACHER INSTRUCTIONAL SKILLS IMPLEMENTATION

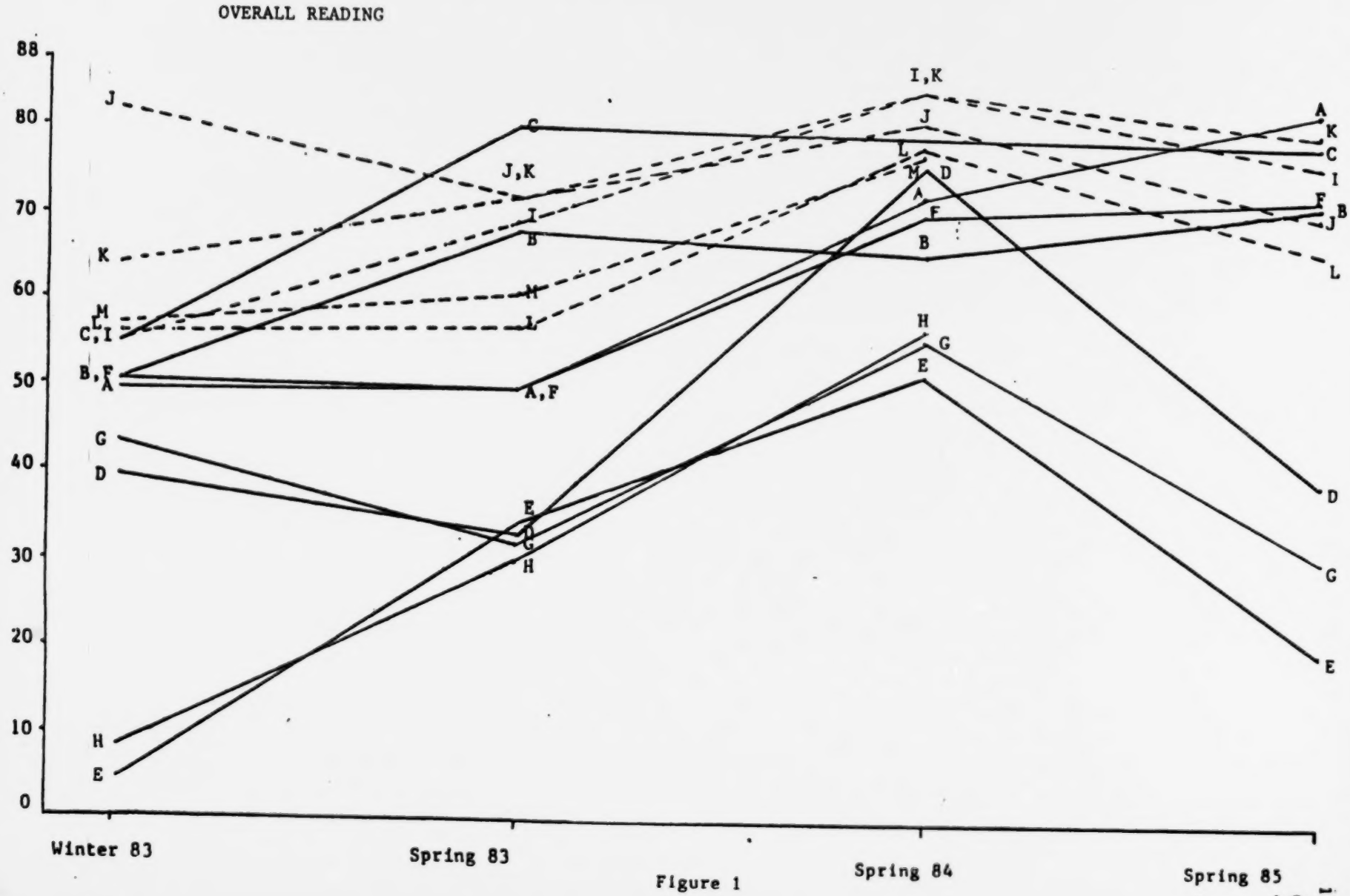
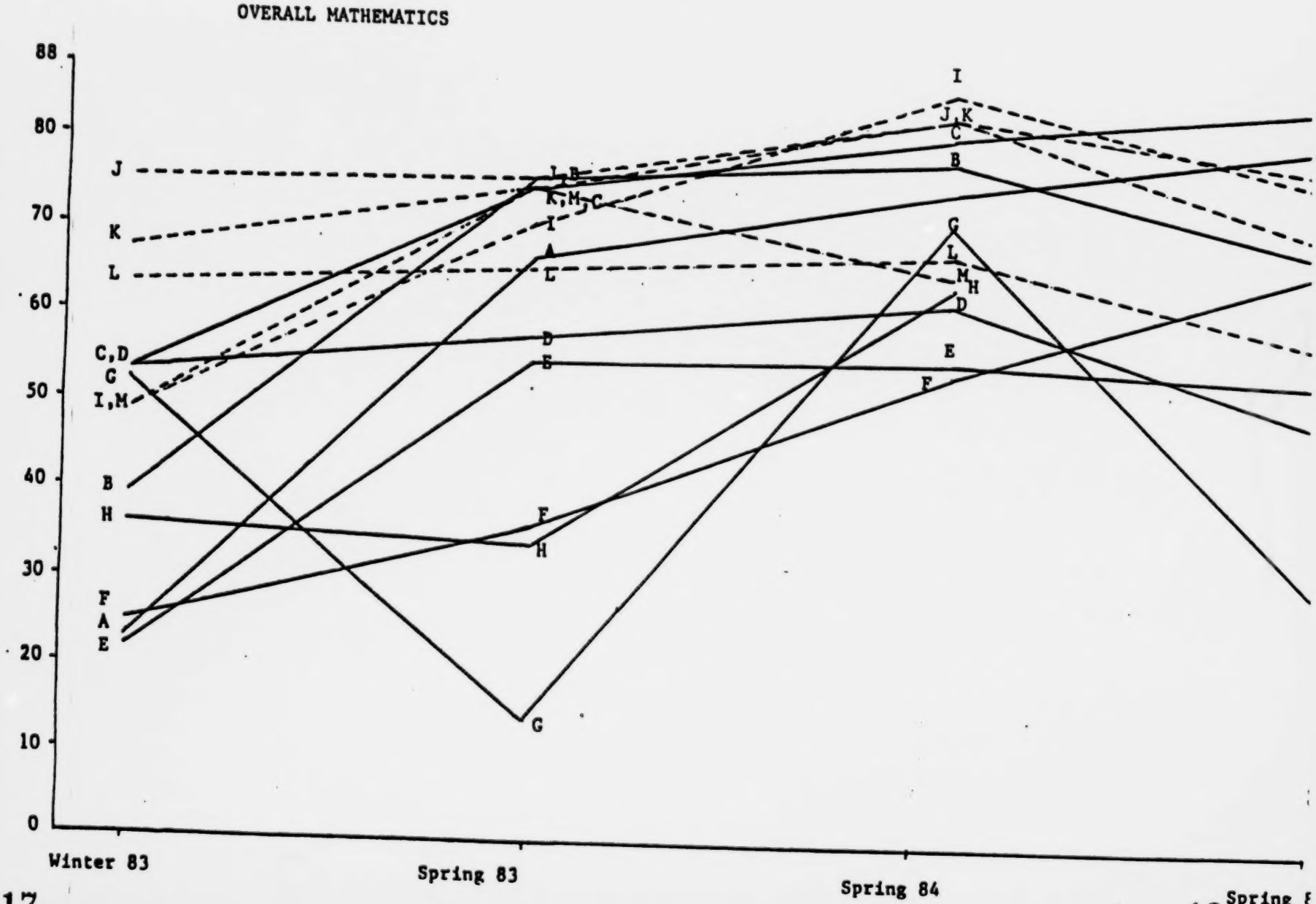


Figure 1

*Raw scores were used in this figure

TEACHER INSTRUCTIONAL SKILLS IMPLEMENTATION



*Raw scores were used in this figure.

Figure 2

Program Impact on Student Engaged Rate

Project level engaged rate. To examine the change in student engaged rate over the course of the project, one-tailed t-tests were computed. Normal curve equivalency scores were used in the analyses. All of the students were used in the analyses who had the required engaged rate scores for the analyses each year. For example, for spring 1983 and spring 1984, there were 192 students with both years' engaged rate scores in reading. In contrast, only 170 students had both 1984 and 1985 engaged rate scores in mathematics.

Overall for the project, the data on Table 3 indicate a significant ($p<.001$) improvement in student engaged rate between spring 1983 and spring 1984. The reverse pattern is found between 1984 and 1985. More specifically, in reading the students were engaged at 84% of the observed time in 1983, 94.6% in 1984, and 88% in 1985. The 1985 engaged rate is higher than when the study started in 1983.

The explanation for the overall drop in engaged rate may again be explained by the late time in the school year when the data were collected. Children and teachers seem to devote less attention to academics as the school year closes, as evidenced by time taken for school photographs, school picnics, parent days, etc.

Engaged rate analyses by grade level and achievement. It is commonly assumed that higher-achieving students are more engaged in their academic work than are low-achieving students. Further, for developmental reasons, older children such as ten-year-old fourth graders might be expected to have developed longer attention spans and, consequently, be involved in academic tasks for longer periods than six-year-old first graders. To examine these hypotheses, an

Table 3

T-tests on Engaged Rates* at Project Level - Spring 1983-Spring 1984 and Spring 1984-Spring 1985

Year	Mean	Difference of Mean	Standard Deviation	T-value	Degree of Freedom	One-Tailed Probability
READING						
Spring 1983	2.327 84%					
		+ .347	.59	8.24	192	.000
Spring 1984	2.674 94.6%					
Spring 1984	2.654 94%					
		- .222	.599	-4.84	170	.000
Spring 1985	2.432 88%					
MATH						
Spring 1983	2.315 83.5%					
		+ .335	.524	9.19	206	.000
Spring 1984	2.649 93.8%					
Spring 1984	2.683 94.9%					
		- .126	.715	-2.36	178	.009
Spring 1985	2.558 91.5%					

*Engaged Rate scores are converted to arcsins.

analysis of covariance was computed using a 4x3 (Achievement Quartile by Grade) factorial design. Achievement was defined by national quartiles. Spring 1984 engaged rates were held constant making it impossible to use 1985 first grade classes in this analysis.

Grade level and engaged rate. A significant main effect of grade was found showing differences in amount of engaged rate between grade levels. These differences may be associated with developmental attention span (see Stallings, Goodman, and Johnson 1985). Fourth graders were engaged approximately 90% of the time in both reading and mathematics (see Table 4). Second graders were engaged 11% less frequently than fourth graders in reading and 14% less frequently in mathematics. However, there was no difference in the amount of achievement gain made at the different grade levels. This may suggest that achievement gain in second grade requires less engaged time than does achievement gain in the fourth grade.

Achievement level and engaged rate. A difference in engaged rate was not found at different achievement levels. In reading, the range of quartile means was from 82.24% to 84.40%. Those students scoring above the 75th quartile in reading achievement had a slightly lower engaged rate than those students scoring between the 26th and 50th quartile in reading achievement.

In mathematics the lowest achievers had the lowest engaged rate and the highest achievers had the highest engaged rates; but, the differences are not statistically significant.

Another interesting question is, "Is the engaged rate of students greater at some achievement levels than at others?" The

Table 4

Program Impact on Engaged Rate
Percent Engaged Rate by Grade by Achievement Quartile
(ANCOVA with 1984 Engaged Rate as Covariate)
Separate Analysis for Reading and Mathematics

READING (Percent Engaged Rate)

Grade	Achievement Quartile 1	2	3	4	
2	74.43 (11)	83.93 (12)	83.33 (10)	75.67 (5)	79.65 (39)
3	82.65 (13)	78.91 (12)	76.31 (14)	79.22 (15)	79.41 (55)
4	91.41 (10)	87.80 (21)	92.91 (13)	95.00 (9)	90.96 (53)
$\bar{X}$	82.24% (N=34)	84.40 (N=45)	84.04 (N=37)	83.50 (N=29)	N=145

*Significant Main Effect of Grade $F(2,132)=8.02$, $p<.01$

Each mean for grade level significantly differs from every other using the Scheffe method for multiple comparisons.

MATHEMATICS (Percent Engaged Rate)

2	73.48 (11)	85.64 (10)	68.99 (11)	76.99 (6)	76.04 (39)
3	91.12 (7)	91.12 (12)	95.08 (20)	92.00 (15)	92.96 (55)
4	90.59 (9)	88.10 (11)	88.88 (21)	94.79 (11)	90.45 (53)
$\bar{X}$	83.76 (N=27)	88.45 (N=33)	87.06 (N=52)	90.25 (N=32)	N=147

*Significant Main Effect of Grade, $F(2,131)=12.82$, $p<.001$

Each mean for grade level significantly differs from every other using the Scheffe method for multiple comparisons.

NOTE: The analysis of covariance was run using arcsine converted engaged rate scores. Means are reported in percents to accommodate interpretation.

analysis of covariance indicated no differential effects of achievement quartiles on 1985 engaged rate scores for reading or mathematics, holding 1984 engaged rate constant (see Table 4).

Continuing students compared with new students. Since this was a longitudinal study, we may ask whether students who had been in the program for more than one year had higher engaged rates than students who entered during the final year of the study. To examine this question, one-tailed t-tests were computed using students who had engaged rate scores for more than one year and students with only 1985 engaged rate scores. The scores of students who were in the program longer were significantly higher than the scores of students new to the program in mathematics but not in reading (see Table 5). The continuing students' engaged rate scores in reading were higher but not significantly.

Engaged rates by teacher classrooms. Another way to think about the impact of the program upon engaged rate is to consider each teachers' impact on different groups of students. The pattern of engaged rate for each teacher's classroom from baseline through the final year of the study is shown on Figures 3 and 4. There are three statistical differences in reading engaged rate scores between 1984 and 1985 for the students of teachers D, B, J, and L. The 1984 students of those teachers were engaged in their work more than were the 1985 students. In mathematics, six out of 11 teachers had students with lower engaged rates in 1985 than in 1984.

Program Impact on Achievement

The most important question to be answered by this study is, "Have the children in the Napa Follow Through program thrived

Table 5

T-tests for Differences in Engaged Rate[x]
Between Continuing vs. New Students
Reading and Math

	N	Mean	Standard Deviation	t-value	Degree of Freedom	One-tailed Probability
READING						
Continuing Students	183	2.46 89%	.54	1.06	250	p=.144
New Students	69	2.37 86%	.62			
MATHEMATICS						
Continuing Students	199	2.57 92%	.61	2.99	268	p=.0015*
New Students	71	2.31 84%	.70			

[x]Engaged rate scores are converted to arcsin.
*Continuing students had higher engaged rate than new students.

Continuing = any student who entered program 1984 or before
New = entered in 1985
N = Number of cases

Figure 3 Mean Reading Engaged Rates

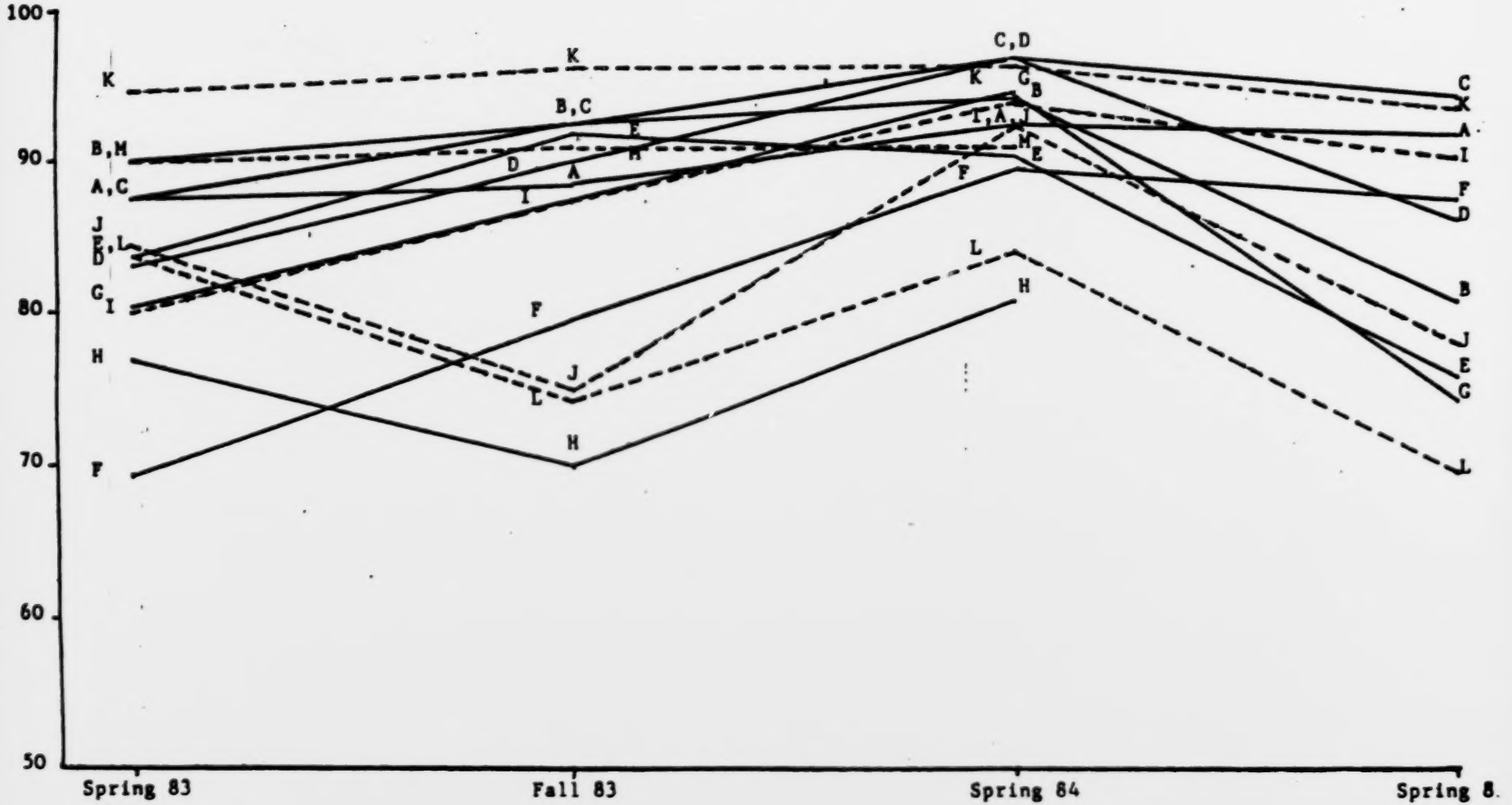
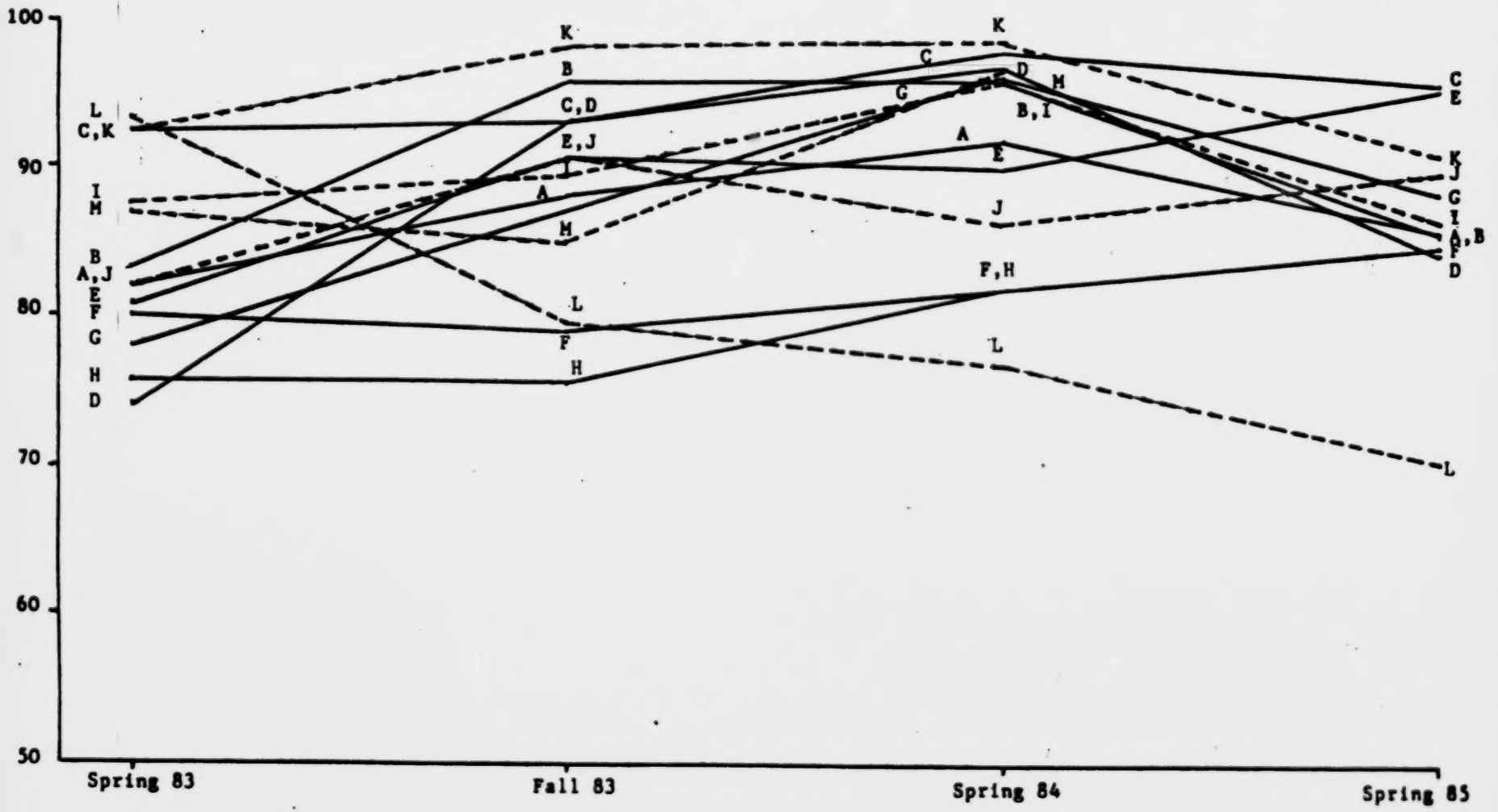


Figure 4 Mean Mathematics Engaged Rate



academically?" Several approaches have been used to answer this question. First, the achievement of children who had been in the project for several years was examined. Did they achieve more than students who were in the program a year or less? Did the program children achieve more than was expected of children in school populations with similar backgrounds? Finally, did the program children achieve more than children in control schools? These questions have been addressed at the project level using individual student scores and classroom scores.

The longitudinal impact. To assess the longitudinal impact of the Napa Follow Through program, t-tests were computed to compare the year-to-year achievement of those students who entered the program in 1982 and stayed until 1985. There were 102 students who stayed in the program all four years (see Table 6). In reading these children gained in the first year, although not statistically, and showed statistical gain in the second year of the project. Mathematics scores increased in the first and second year of the study, although not statistically. In 1985 achievement scores dropped 2.44 NCEs in reading and .18 NCEs in mathematics. The gain for those children remaining in the project from 1982 to 1985 was 1.98 NCEs in reading and 2.28 NCEs in mathematics.

Another way to consider improvement is to see to what degree students moved from the lower half of national test percentiles to the upper half (above the 50th percentile). For the project there was an increase of 12% in students scoring above the 50th percentile in reading and 15% above the 50th percentile in mathematics.

To understand whether the students leaving the program might

Table 6

T-tests for Students Who Stayed in the Program All Four Years (1982-1985)

READING	N of Cases	Mean	Difference of Mean	Standard Deviation	T-value	Degree of Freedom	One-tailed Probability
1982	102	44.45	+1.64	15.99	+1.04	101	.15
1983		46.09					
1983	102	46.09	+2.78	14.87	+1.89	101	.03*
1984		48.87					
1984	102	48.87	-2.44	13.07	-1.89	101	.03*
1985		46.43					
1982	102	44.45	+1.98	17.43	+1.15	101	.13
1985		46.43					
MATHEMATICS							
1982	102	49.60	+1.74	16.18	+1.09	101	.13
1983		51.35					
1983	102	51.35	+.72	16.18	+.45	101	.32
1984		52.07					
1984	102	52.07	-.18	13.59	-.13	101	.45
1985		51.88					
1982	102	49.60	+2.28	19.63	+1.17	101	.12
1985		51.88					

be higher achieving and perhaps from more upwardly mobile families, a comparison was made between the 1982 baseline scores of students who stayed in the project and those who left. This analysis identified a statistical difference ($p<.05$) in favor of the students who left the program some time before 1985.

Continuing students compared to new students. Both project schools seem to attract low income families new to the area. When the parents' work called them to another area or they could afford to move to a better neighborhood, the families moved. Given the revolving doors of these schools, it is of interest to compare the continuing students with the new students. New students were defined as those entering school during school year 1984-1985; continuing students were those who entered the program in 1982-1984. T-tests were computed to test the differences in continuing and new students in 1985 achievement. There were no significant differences between the two groups in either reading or mathematics, although continuing students did score higher in both subjects (see Table 7).

Differential effects of grade level and achievement level. It is possible for educational programs to effect children differently at different grade levels, and programs might effect high, average, or low achieving children differently. Who thrives is an important question to ask of any school intervention. To learn how the Follow Through program effected different groups of children, an analysis of variance was computed holding 1984 achievement scores constant. A 4x3 factorial design included four achievement levels (defined by national quartiles) and three grade levels. Since the 1985 first graders did not have 1984 achievement

Table 7

Continuing vs. New Students
T-tests on Achievement 1985

	N of Cases	Mean	Standard Deviation	T-value	Degree of Freedom	One-tailed Probability
READING						
Continuing Students	231	46.91	19.56	.68	297	.248
New Students	68	45.12	17.64			
MATHEMATICS						
Continuing Students	232	50.78	21.13	.66	296	.253
New Students	66	48.86	19.39			

No significant differences.

scores, they could not be included in these analyses (see Table 8).

Results for both the reading and mathematics analyses indicated only the expected significant difference between quartile levels. There were no achievement differences between grade levels, holding 1984 achievement constant, and no quartile level by grade interactions.

Expectancy Bands for Achievement

In the state of California tests have been developed to examine elementary students in reading and mathematics in third and sixth grade each year. The California Assessment Program (CAP) tests are highly correlated with the Comprehension Test of Basic Skills and the Stanford Achievement Test which are also administered by the school districts of Treatment Schools #1 and #2. CAP tests are administered yearly to all third and sixth grade students and are of interest to this study since the state CAP and local standardized test results can be compared.

The California Assessment Program (CAP) has established an expectancy score band of achievement performance from a battery of state tests for each school in the state. The expectancy score bands are calculated using three background factors: Socioeconomic Status (SES), Aid to Families with Dependent Children (AFDC), and limited English speaking (LES). Schools serving similar populations would be expected to perform within the same expectancy score band. For example, if a school's reading score falls within the comparison score band, the school's reading program may be considered typical of the programs in similar schools in the state. A score above the range of the band may indicate a program better than the programs typical of similar schools in California. The

Table 8

Program Impact on Academic Achievement
Mean Achievement (NCE) by Grade and by Achievement Quartile
(ANCOVA with 1984 Achievement Scores as Covariate)
Separate Analyses for Reading and Mathematics

Grade	Achievement Quartile 1	2	3	4	
READING					
2	25.47 (N=15)	42.99 (N=16)	57.60.(N=12)	73.45 (N=7)	45.51 (N=50)
3	18.60 (N=21)	45.15 (N=15)	57.77 (N=15)	75.12 (N=14)	45.93 (N=65)
4	26.05 (N=12)	43.51 (N=28)	58.79 (N=21)	72.27 (N=12)	49.76 (N=73)
$\bar{X}$	22.61 (N=48)	43.79 (N=59)	58.17 (N=48)	73.72 (N=33)	N=188

*Significant Main Effect of Quartile $F(3,175)=147.26$ $P<.0001$

Each group (Achievement Quartile) is significantly different from each other group, using a Scheffe Multiple Comparison.

MATHEMATICS					
2	24.09 (N=16)	44.82 (N=9)	56.36 (N=15)	76.75 (N=10)	47.67 (N=50)
3	20.34 (N=11)	41.52 (N=18)	56.21 (N=21)	74.66 (N=15)	50.33 (N=65)
4	29.08 (N=12)	41.09 (N=16)	56.47 (N=25)	78.78 (N=20)	54.71 (N=73)
$\bar{X}$	24.09 (N=39)	42.05 (N=43)	56.36 (N=61)	76.95 (N=45)	N=188

*Significant Main Effect of Quartile $F(3,175)=145.74$, $P<.0001$

Each group (Achievement Quartile) is significantly different from each other group, using a Scheffe Multiple Comparison.

parameters of the expectancy band may change from year to year if the background factors of the school districts change.

Using the CAP data, trends for the treatment and control schools are examined from 1981 to 1985. All third and sixth grade students were tested each year. These data are reported on Table 9.

It is of interest to see whether schools performed above, below, or within the expected score band. The students in third and sixth grade classrooms of Treatment School #1 were above their expected score band in reading and mathematics every year except 1984-1985: the third grade scores dropped 23 points in reading and 16 points in mathematics. The scores were lower in that year than in any other year. Control school #1's students tested within their expectancy band on 11 of 16 tests reported. In 1984-1985 the third grade control students' reading scores were similar to the previous years but gained 35 points in mathematics.

The drop in Treatment School #1's CAP scores in 1984-1985 parallels a drop in the national standardized achievement scores. The decrease in scores cannot be attributed to a shift in student background factors since the school's SES rank was considerably higher in 1984-1985 than in 1983-1984.

In Treatment School #2 the third graders gained 44 points in reading between 1982 and 1983 and then scored about the same during the next two years of the study. In 1985 the third graders scored 17 points higher in mathematics than in 1984: this was above expectation. From the inception of the program in 1981, third grade mathematics scores went up 38 points on the CAP. Over a similar time period the control school scores went up 5 points in

Table 9

Achievement Expectancy Bands for
Treatment and Control Schools
Reading and Mathematics

Schools	Year	N	Reading Score	Reading EB Low-High	Mathematics Score	Mathematics EB Low-High
#1 Treatment						
Grade 3	1981-1982	52	268(A)	228 266	341(A)	222 276
	1982-1983	56	264(A)	215 253	283(A)	222 260
	1983-1984	69	271(A)	215 253	277(A)	225 260
	1984-1985	69	248(W)	238 269	261(W)	240 273
Grade 6	1981-1982	57	258(A)	220 254	267(A)	226 259
	1982-1983	61	259(A)	220 253	260(W)	228 261
	1983-1984	66	254(A)	201 233	268(A)	214 247
	1984-1985	44	249(W)	249 284	244(W)	255 292
#1 Control						
Grade 3	1981-1982	56	254(W)	241 275	259(W)	241 278
	1982-1983	54	294(A)	239 280	275(W)	241 279
	1983-1984	55	286(W)	258 297	252(B)	262 295
	1984-1985	54	285(W)	246 287	287(W)	247 288
Grade 6	1981-1982	63	269(A)	234 266	279(A)	239 271
	1982-1983	61	252(W)	243 274	265(W)	248 280
	1983-1984	66	254(W)	239 272	256(W)	250 281
	1984-1985	55	291(A)	226 264	271(W)	238 272
#2 Treatment						
Grade 3	1981-1982	25	217(W)	217 258	245(W)	222 267
	1982-1983	31	261(W)	214 261	287(A)	217 271
	1983-1984	32	258(W)	217 261	266(W)	222 270
	1984-1985	23	259(W)	218 265	283(A)	222 278
Grade 6	1981-1982	39	252(A)	214 250	268(A)	219 258
	1982-1983	30	270(A)	213 258	281(A)	223 269
	1983-1984	25	263(A)	195 242	275(A)	208 260
	1984-1985	23	224(W)	211 254	279(A)	219 273
#2 Control						
Grade 3	1981-1982	71	319(A)	248 280	281(W)	245 281
	1982-1983	68	264(W)	246 281	292(A)	244 282
	1983-1984	64	303(A)	242 281	300(A)	244 286
	1984-1985	75	289(W)	254 291	286(W)	255 292
Grade 6	1981-1982	111	282(A)	241 268	269(W)	246 271
	1982-1983	96	267(W)	245 272	268(W)	251 277
	1983-1984	84	272(A)	238 270	282(A)	250 279
	1984-1985	79	265(W)	238 268	277(W)	248 277

N = Number of students tested
Score - Scaled Scores (A=above expectancy; W-within expectancy; B=below expectancy)
EB = Expectancy Band

mathematics. Sixth grade students tested above the expected band for mathematics in all years of the study. In reading the sixth graders scored above what was expected each year except 1984-1985.

The control school for School #2 scored above their expectancy band in two of the four years investigated in both reading and mathematics. However, from 1982 to 1985 third grade reading scores dropped 30 points, sixth grade reading scores lost 17 points. Both grades had higher mathematics scores in 1985 than in 1982.

Clearly, Treatment School #2 third grade students made more progress in reading and mathematics than did the control third graders. The third grade students remaining in the study would have entered the program as first graders. This may be one indication of a program accumulative effect. An alternative explanation might be the outstanding quality of the third grade teacher in School #2. This teacher scored very high on the ISOI.

Control School Comparisons on Standardized Achievement Tests

The comparison of the amount of gain achieved by treatment and control students is of considerable interest. Table 10 displays the 1984 to 1985 reading and mathematics gain scores for treatment and control students in grades 2 through 4. Treatment students' reading scores in 1985 were less good for that grade than they were in 1984. The decline in achievement was especially dramatic for second grade treatment children. The only treatment gain found between 1984 and 1985 was in third grade mathematics. This finding is similar to the gain reported on the CAP scores for third graders in mathematics.

To statistically compare 1985 achievement of treatment and control students, a 2x3x4 analysis of covariance was performed.

Table 10

Reading Achievement Gain
1984-1985
Treatment Vs. Control

Grade	Treatment	Control	Total
2	-7.86 (N=52)	-3.15 (N=98)	-4.78 (N=150)
3	-.09 (N=67)	+3.01 (N=86)	+1.65 (N=153)
4	-.68 (N=75)	+3.08 (N=90)	+1.37 (N=165)
	$\bar{X}$=-2.40 (N=194)	$\bar{X}$=+.83 (N=274)	

Math Achievement Gain
1984-1985
Treatment Vs. Control

Grade	Treatment	Control	Total
2	-5.91 (N=50)	+2.35 (N=98)	-.44 (N=148)
3	+2.06 (N=65)	-3.69 (N=86)	-1.21 (N=151)
4	-2.22 (N=73)	+3.76 (N=89)	+1.06 (N=162)
	$\bar{X}$=-1.72 (N=188)	$\bar{X}$=+.91 (N=273)	

The dependent variable was 1985 achievement scores, and separate analyses were run for reading and mathematics. Achievement in 1984 was held constant, in order to control for preexisting differences between treatment and control students. Therefore, no first graders could be used in the analysis. Control students outnumbered treatment students because Control School #2 was twice as large as Treatment School #2.

The three independent factors were treatment vs. control group membership, grade levels 2 through 4, and achievement levels.

The reading analyses revealed three main effects and one interaction effect (see Table 11). Control students (N=274) scored significantly higher (8.66 NCE points) than treatment students (N=194) in reading. Achievement scores also significantly differed by grade and by quartile. A quartile by group interaction was found. Treatment students in quartiles 1, 2, and 4 scored lower than control students, while treatment students in quartile 3 scored higher than control students in that quartile. Although statistically significant, this difference is not of practical significance to us since NCE point differences are minimal.

The analysis of mathematics achievement revealed two significant effects (see Table 12). The main effect of quartile is an expected effect, due to the rules that determine group membership. A quartile by group interaction was found for mathematics, as for reading. Here, the effect was found because fourth quartile treatment students outscored fourth quartile control students. For the three other quartiles, control students scored higher. Again, this result is not of any practical significance. The major finding for mathematics was that the

Table 11

Reading Achievement (1985)
(Mean NCE Score)
Treatment vs. Control Comparison
2x3x4 ANCOVA, Holding 1984 Achievement Constant

Grade	Quartile	Treatment	Control	Total
2	1	25.79 (N=16)	28.45 (N=15)	$\bar{X}$=50.57
	2	42.99 (N=16)	45.05 (N=25)	N=150
	3	57.60 (N=12)	56.32 (N=36)	
	4	74.05 (N=8)	73.68 (N=22)	
3	1	18.60 (N=21)	25.66 (N=8)	$\bar{X}$=51.67
	2	45.14 (N=15)	44.18 (N=21)	N=153
	3	57.77 (N=15)	56.59 (N=31)	
	4	75.60 (N=16)	72.07 (N=26)	
4	1	26.07 (N=13)	31.13 (N=6)	$\bar{X}$=55.79
	2	43.51 (N=28)	45.94 (N=22)	N=165
	3	58.79 (N=21)	57.90 (N=25)	
	4	71.86 (N=13)	76.58 (N=37)	
	*Total	$\bar{X}$=47.70 (N=194)	$\bar{X}$=56.36 (N=274)	N=468

Significant Effects

3 Main Effects

Quartile, $F(3,467) = 328.46$, $p<.001$
Grade, $F(2,467) = 4.12$, $p<.02$
Group, $F(1,467) = 450$, $p<.04$

Interaction

Quartile by Group, $F(3,467) = 3.37$, $p<.02$

Table 11 (continued)

Quartile by Group Interaction

Quartile	Treatment	Control	Total
1	22.84 (N=50)	28.24 (N=29)	24.82 (N=79)
2	43.79 (N=59)	45.07 (N=68)	44.47 (N=127)
3	58.17 (N=48)	56.84 (N=92)	57.30 (N=140)
4	73.95 (N=37)	74.45 (N=85)	74.30 (N=122)

Grade by Group Interaction (non-significant)

Quartile	Treatment	Control	Main Effect of Grade (total)
2	45.85 (N=52)	53.08 (N=98)	50.57 (N=150)
3	46.92 (N=67)	55.36 (N=86)	51.66 (N=153)
4	49.68 (N=75)	60.87 (N=90)	55.79 (N=165)

Main effect of grade: Using Scheffe

2nd and 4th grades are significantly different from one another, F=11.01. Other grades are not significantly different.

Main effect of quartile: Using Scheffe

Each quartile significantly differs from every other quartile.

Table 12

Mathematics Achievement (1985)
(Mean NCE Score)
Treatment vs. Control Comparison
2x3x4 ANCOVA, Holding 1984 Achievement Constant

Grade	Quartile	Treatment	Control	Total
2	1	22.92 (N=16)	28.42 (N=12)	$\bar{X}$=53.16
	2	44.60 (N=9)	44.82 (N=27)	(N=148)
	3	56.39 (N=15)	54.31 (N=20)	
	4	76.75 (N=10)	73.18 (N=39)	
3	1	20.34 (N=11)	27.23 (N=12)	$\bar{X}$=52.66
	2	41.52 (N=18)	45.01 (N=20)	(N=151)
	3	56.21 (N=21)	56.98 (N=34)	
	4	74.66 (N=15)	75.81 (N=20)	
4	1	29.97 (N=18)	29.08 (N=12)	$\bar{X}$=59.98
	2	41.09 (N=16)	45.75 (N=10)	(N=162)
	3	56.48 (N=25)	58.15 (N=25)	
	4	78.78 (N=20)	75.42 (N=48)	
		$\bar{X}$=51.32 (N=188)	$\bar{X}$=58.19 (N=273)	

Group Main Effect is not significant.

Significant Effects:

Main Effect Quartile $F(3,460)=485.64$, $p<.001$

Interaction Quartile by Group $F(3,460)=6.02$, $p<.01$

Table 12 (continued)

Quartile by Group Interaction

Quartile	Treatment	Control	Total
1	24.09 (N=39)	28.61 (N=30)	26.06 (N=69)
2	42.05 (N=43)	44.91 (N=57)	43.68 (N=100)
3	56.36 (N=61)	56.68 (N=79)	56.54 (N=140)
4	76.95 (N=45)	74.68 (N=107)	75.35 (N=152)

Grade by Group Interaction

Quartile		
2	47.67 (N=50)	55.95 (N=98)
3	50.33 (N=65)	54.42 (N=86)
4	54.71 (N=73)	64.29 (N=89)

*Using Scheffee, each quartile is different from each other one.

treatment (N=188) and control (N=273) students were not significantly different in achievement, when 1984 scores were controlled.

Relationship Between Engaged Rate and Achievement

The relationship between achievement and engaged rate is of key interest to this project. Using the 1985 students with both achievement test scores and engaged rates, Pearson product moment correlations were computed. A positive correlation ($p<.001$) was found only between engaged rate and mathematics achievement test scores. No relationship was found between reading achievement and engaged rate (see Table 13). A scatterplot of 1985 reading indicates that many students were highly engaged in the task of reading but they were not achieving (see Figure 5). In both 1983 and 1984 significant positive relationships were found between engaged rate and both reading and mathematics achievement.

Although it is commonly assumed that high achieving students have higher engaged rates, the previous analyses indicate this is not always true. To further explore this question, one-way analysis of variance was prepared using quartile achievement levels as the independent variable and engaged rates as the dependent variable (see Table 14). There was no significant difference between quartiles for reading, while there was a significant difference for mathematics. Students in the third and fourth quartile were more engaged than those in the lower two quartiles.

If engaged rate increases, does achievement increase? To address this question, a 2x2 table was compiled comparing increases and decreases in achievement groups with increases and decreases in engaged rate groups. Those at exactly the center were placed in

Table 13

Correlations Between Student Achievement and Student Engaged Rate

	Engaged Rate
1985	
Mathematics Achievement	r = .26* p = .001 (N=333 students)
Reading Achievement	r = .02 p = .35 (N=312 students)
1984	
Mathematics Achievement	r = .14* p = .01 (N=335 students)
Reading Achievement	r = .20* p = .001 (N=340 students)
1983	
Mathematics Achievement	r = .13* p = .01 (N=344 students)
Reading Achievement	r = .12* p = .01 (N=334 students)

*Significant Correlations

Figure 5

Relationship Between 1985 Achievement
and 1985 Engaged Rate (N=312)

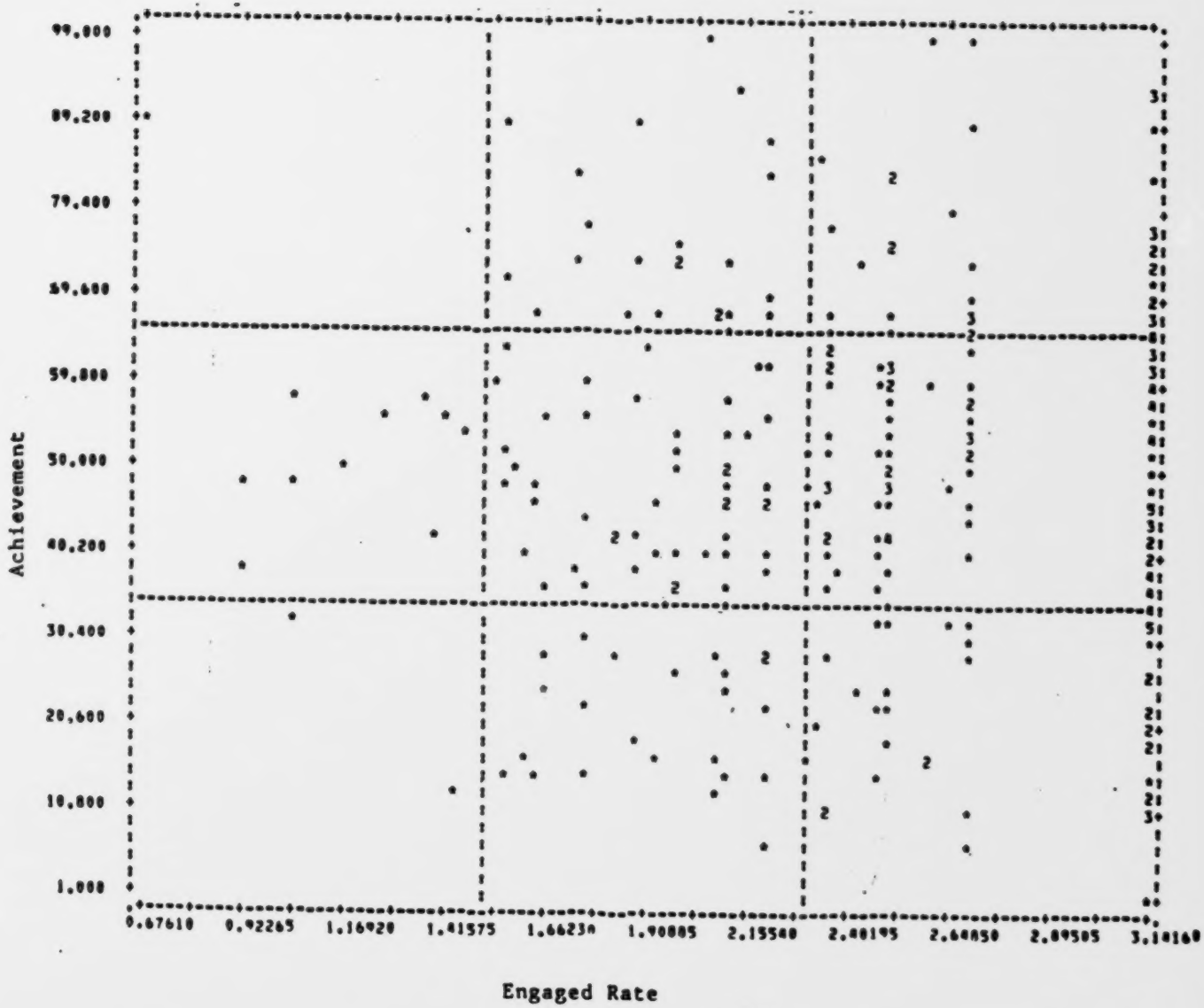


Table 14

One-Way ANOVA on 1985 Engaged Rate
by Achievement Quartile
(Reading and Mathematics)

Reading Achievement Quartile 1985	Mean Reading Engaged Rate 1985
1	2.48 (N=77) 84.1%
2	2.40 (N=86) 81.6%
3	2.48 (N=81) 83.8%
4	2.52 (N=73) 84.8%

no significant differences

Mathematics Achievement Quartile 1985	Mean Mathematics Engaged Rate 1985
1	2.32 (N=69) 81.4%
2	2.45 (N=73) 81.6%
3	2.52 (N=101) 86.0%
4	2.67 (N=95) 90.2%

*Significant Main Effect of Achievement Quartile
$F(3,337)=4.65$, $p<.01$

One-way analysis of variance

Dependent variable and engaged rate

Independent variable and achievement quartile

the increase group. The differences in distribution of students in the groups were analyzed using chi-square tests. For reading the groups were distributed in a way that was not significantly different from chance. For mathematics, however, the chi-square was significant.

In reading 42% of the students decreased their engaged rate and decreased their achievement from 1984 to 1985. Only 17% both increased their reading scores and increased their engaged rates. In mathematics the results were similar (see Table 15).

The finding regarding the percent of children who decreased their engaged rate and achievement is particularly alarming. One possibility was that children with decreasing scores could be found in a few classrooms where teachers were implementing the program to a low degree. However, decreasing students were found to some degree in all classrooms. The highest percentage was in two new teachers' classrooms and that of teacher G whose ISOI patterns were so erratic (see Table 16).

Program Impact on LES Students

Included in the student population of the Napa project were approximately 30% limited English speaking students (LES). Of interest is how these children have fared during the life of the project. Our t-test analysis compared achievement and engaged rate of LES students with English speaking students from 1982 to 1985. The engaged rates comparing LES students and their English speaking counterparts indicated significantly higher engaged rates for the LES students in 1984 and 1985 reading. At the other data points there were no differences (see Table 17). As might be expected, the results of these analyses indicated that the English speaking

Table 15

Relationship Between Change in Achievement and Change in Engaged Rate

	Increased Achievement 1984-1985	Decreased Achievement 1984-1985	
READING			
Increased Engaged Rate 1984-1985	Frequency=17% (N=25)	Frequency=19% (N=29)	36% (N=54)
Decreased Engaged Rate 1984-1985	Frequency=22% (N=34)	Frequency=42% (N=63)	64% (N=97)
	39% (N=59)	61% (N=92)	
	$x^2(1)=1.40$, N.S.		
MATHEMATICS			
Increased Engaged Rate 1984-1985	Frequency=21% (N=31)	Frequency=22% (N=33)	43% (N=64)
Decreased Engaged Rate 1984-1985	Frequency=15% (N=22)	Frequency=42% (N=64)	57% (N=86)
	36% (N=53)	64% (N=97)	
	$x^2(1)=7.42$*, p=.006		

Table 16

Students Who Decreased From 1984 to 1985 in Both Achievement and Engaged Rate

READING		
Teacher	#of Students	% of Students
N	6	9.5
B	1	1.6
C	2	3.2
D	6	9.5
E	3	4.8
F	8	12.7
G	11	17.5
O	4	6.3
P	2	3.2
J	3	4.8
K	3	4.8
L	7	11.1
Q	7	11.1

MATHEMATICS		
Teacher	#of Students	% of Students
N	8	12.7
C	1	1.6
D	3	4.8
E	1	1.6
F	1	11.1
G	8	12.7
O	8	12.7
J	3	4.7
K	1	1.6
L	12	18.8
Q	11	17.2

Table 17

T-tests on Engaged Rates in Reading and Mathematics:
English Speaking vs. Limited English Speaking for 1983-1985

READING

Group	Number of Students	Mean	Standard Deviation	T-value	Degrees of Freedom	One-tailed Probabili
Spring 1983 English Speaking	161	2.35	.44	1.92	188	.056
Spring 1983 Limited English Speaking	29	2.18	.48			
Spring 1984 English Speaking	160	2.63	.36	-1.64	189	*.051
Spring 1984 Limited English Speaking	31	2.71	.38			
Spring 1985 English Speaking	134	2.37	.54	-1.93	195	.026*
Spring 1985 Limited English Speaking	63	2.56	.57			

MATH

Spring 1983 English Speaking	167	2.30	.42	-.74	198	.23
Spring 1983 Limited English Speaking	33	2.34	.40			
Spring 1984 English Speaking	164	2.60	.41	-.73	193	.23
Spring 1984 Limited English Speaking	31	2.66	.39			
Spring 1985 English Speaking	137	2.54	.60	- .49	208	.312
Spring 1985 Limited English Speaking	73	2.58	.76			

children scored higher on reading and mathematics achievement tests than did LES students. However, more achievement gain was made in both reading and mathematics by the LES students each year of the study (see Table 18). From 1982 to 1985 the LES students raised their reading scores 10.98 NCE points and their mathematics scores 6.71 points. Both LES and English speaking students' scores dropped in reading in 1985 but the drop was far more severe for the English speaking students.

The principal at the school which served the LES children tested indicated that LES students in the project were not pulled from class but were taught in the classroom with the help of a bilingual aide where necessary. This school used an immersion process so that children received all of their lessons in English to the greatest degree possible. The district had developed their own materials and program to help LES students adapt as quickly as possible to the school routines and expectations.

The Relationship Among ISOI Variables, Engaged Rate, and Achievement

As previously stated, the four composite instructional variables (set, instruction, guided practice, and independent practice) were highly intercorrelated in both reading and mathematics. This indicates similar dimensions were being measured by these variables. Only instruction and independent practice were not significantly related in mathematics. Even though the four variables may be measuring similar dimensions, it is of interest to understand the relationships of the separate variables with student outcomes in reading and mathematics.

Table 18

T-tests Between Limited English Speaking and English Speaking Students
Reading and Math Achievement, 1982 through 1985

Group	Degree of Freedom	Mean	Standard Deviation	Pooled Variance Estimate T-value	Pooled Variance Estimate One-tailed Probability
READING					
1982 English Speaking	186	50.02	84.10		
1982 Limited English Speaking	30	23.63	13.68	2.00	*.02
1983 English Speaking	186	49.73	20.94		
1983 Limited English Speaking	35	26.25	16.67	6.08	*.000
1984 English Speaking	170	52.81	17.77		
1984 Limited English Speaking	30	35.25	15.25	7.33	*.000
1985 English Speaking	170	50.84	18.39		
1985 Limited English Speaking	73	34.61	18.32	6.35	*.000
MATH					
1982 English Speaking	186	52.83	21.05		
1982 Limited English Speaking	30	32.78	17.19	4.79	*.000
1983 English Speaking	186	53.04	19.98		
1983 Limited English Speaking	35	38.92	15.15	4.02	*.000
1984 English Speaking	159	53.39	17.94		
1984 Limited English Speaking	35	42.17	14.89	4.62	*.000
1985 English Speaking	171	54.02	20.91		
1985 Limited English Speaking	73	39.49	17.72	5.22	*.000

In previous analyses, (1983 and 1984), the correlations between total ISOI, engaged rate, and achievement have not been consistent. To examine these relationships in greater depth, Pearson product moment correlations were computed using 14 classrooms (including the teachers new to the project) and the students with the required outcome scores.

ISOI and engaged rate. The results of these analyses indicated significant relationships between reading engaged rates and all components of the ISOI except independent practice (see Table 19). During independent practice students were expected to work alone. Apparently, these students had difficulty staying engaged in seatwork during independent practice.

In mathematics the only significant relationship identified was between engaged rate and guided practice. In guided practice teachers were checking students' understanding of the lesson and conducting some drill and practice to reinforce learning. The goal of this activity was to have all students participating and involved. Apparently, this instructional strategy was effective in helping students stay engaged in their mathematics lessons.

ISOI and achievement. In 1985 a significant relationship was found between student achievement and reading instruction but not with the other components (see Table 20). The instruction component included explanations, student participation (e.g. reading aloud), checks for understanding, feedback, monitoring student progress, and adjusting the lesson to the correct level of difficulty. The lessons were supposed to be appropriate for the students: not too hard and not too easy. All of the 14 teachers divided their students into three or more small groups for reading,

Table 19

Relationship Between Instructional Behaviors and Engaged Rate

	Anticipatory Set	Instruction	Guided Practice	Independent Practice[x]	Total
Mean Student READING Engaged Rate	r=.55* (N=14) p=.02	r=.52* (N=14) p=.02	r=.50* (N=14) p=.03	r=.41 (N=13) p=.08	r=.35 N=14 p=.11
Mean Student MATHEMATICS Engaged Rate	r=.27 (N=14) p=.176	r=.19 (N=14) p=.26	r=.48* (N=14) p=.04	r=.28 (N=13) p=.18	r=.54 N=14 p=.02

*Significant at $p < .05$

[x]One first grade teacher did not use independent practice.

Table 20

Correlations between Teacher Behaviors and Student Achievement 1985

	Anticipatory Set	Instruction	Guided Practice	Independent Practice	ISOI Total
READING					
Mean Student Achievement 1985	r=.11 (N=14) p=.35	r=.52* (N=14) p=.03	r=.38 (N=14) p=.09	r=.39 (N=13) p=.09	r=.33 N=14 p=.12
MATHEMATICS					
Mean Student Achievement 1985	r=.06 (N=14) p=.42	r=.07 (N=14) p=.40	r=.59* (N=13) p=.01	r=.33 (N=14) p=.17	r=.28 N=14 p=.17

and consequently a large amount of the teacher's time was spent instructing as defined by the ISOI. When students were not being instructed by the teacher, they usually worked alone as defined by independent practice. As previously stated, little association was found between independent practice and engaged rate. The challenge seems to be to develop strategies to keep students involved when the teacher is not immediately available.

The only significant relationship between ISOI and mathematics achievement was guided practice. This finding was similar to the ISOI and engaged rate findings. The guided practice component required the active participation of all students and was most likely to keep them involved. The teacher checked all students' understanding and retaught as necessary. This checking and rechecking continued until the teacher was confident the students could work independently. Apparently for teaching mathematics, guided practice was an important strategy.

SUMMARY AND INTERPRETATION

The primary impact of the program culminated at the end of the second year of implementation in spring, 1984. At this time, the teachers' ISOI scores and the students' engaged rates and achievement test scores were at their peak and the relationships among these variables was the most strong. In 1985 the teacher observed behaviors on the ISOI dropped considerably. Student engaged rate decreased, but most alarming was the drop in students' reading test scores. Reading test scores on the nationally standardized CAT or SAT as well as the state CAP, were considerably lower in the 1985 testing compared to 1984 testing. However, the drop was not to the level of scores at the program's inception. A

drop in mathematics achievement test scores was evident on the nationally standardized tests in 1985 in all grade levels. On the state CAP mathematics tests, however, which tested third and sixth grades, a substantial gain was found in mathematics from 1982 to 1985 in the third grade at School #2.

One explanation for the overall decrease in scores may be a regression toward the mean. The teachers and students had received a great deal of attention during 1981-1984 and all scores were high. Perhaps this high level could not be maintained. The downward trend in all evaluation scores was discussed among the project directors, trainers, teachers, and principals. Several explanations were offered. First there was a review of the ISOI components in 1984 just prior to the ISOI observations. This review may have inflated the ISOI scores in 1984. Such a review did not occur in 1985. Certainly the level of project staff involvement was much reduced in 1985 from the previous years. During the final year teachers attended only two project staff development events. There were fewer days allowed for release time from classes to attend special sessions. The final year of the study was focused upon helping the teachers and principals maintain their momentum and start sharing the instructional skills they had gained with teachers in other schools. In fact, several of the project teachers became Mentor Teachers in the California state teacher incentive system. Other teachers have taken a leadership role in providing staff development for the district. However, some teachers reported feeling tired of being in a show case for so long. They loved the project and the staff, but were tired and a little bored with the same ideas after three years. They were

ready to relax and then move on to something else.

Those teachers who had integrated the strategies into their daily planning used them easily on a regular basis. Two teachers who resisted the program initially were still resisting it at the end. It may be a waste of resources to have reluctant teachers in this type of intensive program with such comprehensive opportunities to improve instruction.

The parallel drop in the program's student engaged rate and achievement may be related to some reduced effort on the part of the teachers. Engaged rate continued to relate to instruction in reading and guided practice in mathematics. Independent seatwork continued to be related to student off-task rate. Nevertheless, overall engaged rates were lower in 1985. It is possible that collecting the engaged rate data in late May might have caused a higher off-task rate. The schedule of data collection does not explain the drop in achievement scores since the achievement tests were administered at the usual time of the year under standard conditions. The drop in SAT and CAP test scores in 1985 was hard to understand. One consideration might be a change in the type of population being served at the schools. However, School #1 had a higher SES level in 1985 while School #2 had a lower SES level, and yet the drop in achievement scores was similar for both schools. Both schools lost approximately 30% of their students each year. Importantly, those students continuing in the study for more than one year did score higher than students new to the program.

This evaluation reveals peaks and valleys and hopefully plateaus from which to grow. Over the course of the study, from 1982 to 1985, reading achievement increased 1.98 NCE points and

mathematics achievement increased 2.28 NCE points. However, these increased scores were not greater than the increases for control schools. Certainly the control schools were unusually good. One control school had a very vigorous program of positive reinforcement for students and teachers. This principal had instituted a well organized behavior modification system where students' personal conduct and academic achievements were rewarded daily, weekly, and monthly with school pencils, and treats from fast-food chains and ice cream parlors. Rewards were for such things as showing up on time, completed lessons, keeping the campus clean, helping other people, staying on task. Teachers in both control schools had received from their districts a staff development program which included instruction similar to that of the Follow Through program. The training was not as extensive nor did it have the type of follow-up that occurred in the Follow Through project schools. However, the fact is, the control schools were not standing still.

The good news is that many students in both treatment and control schools were operating at a higher level than was expected for students with their background factors. Students in the treatment program have made progress in all achievement quartiles. Twelve percent more students were scoring above the 50th percentile in reading and 15% more in mathematics than before the program began. Students who had been in the program for longer than one year had higher test scores than those who had been in the program for just one year. Third grade students made gains in mathematics on both the CAP and standardized tests. The teacher of these students warrants a case study to examine what was occurring that

was different from other teachers. Especially good news is the progress the limited English speaking students have been making in the program.

Several important questions have been raised regarding staff development. How much time is needed to implement a staff development program? What does it take to maintain it? These teachers received a great deal of release time for training sessions, 10 days during the first two years of the project (1982-1983 and 1983-1984). They received regular coaching, modeling and feedback from the Follow Through staff for the first two years. The third year was meant to be a year for the principal and faculty to maintain the instructional program with little help from the Follow Through project staff. The project staff provided information, gave assistance, and provided feedback to the teachers upon request but provided little formal instruction. The teachers in interviews with evaluators projected self-esteem and a sense of professionalism. It may be that as in most learning, there are peaks and leveling off periods. Perhaps these teachers will catch a second breath on this plateau and continue to use the instructional methods they have been learning over the past three years. Given the results of the first two years, renewed enthusiasm may well lead to higher student engaged rates and more gain in reading and mathematics. The important question remains. Will the students' 1985-1986 test scores remain at the present level, will the scores decline back to the 1982-1983 level, or will the scores resume an upward trend? Such questions regarding this and other staff development effects are needed to answer questions of cost effectiveness.

REFERENCES

Hunter, M. (1976) Prescription for Improved Instruction. TIP Publications, P. O. Box 514, El Segundo, California 90245.

Stallings, J., Goodman, J., and Johnson, R. "Engaged Rates: Does Grade Level Make A Difference?", (In Press) Journal of Research in Childhood Education.

Stallings, J. (1985) A Study of Implementation of Madeline Hunter's Model and Its Effects on Students. Journal of Educational Research. July/August 1985 (Vol. 78, No. 76), pp. 325-337.

Stallings, J. with Pamela Robbins, Laurie Presbrey, and Jane Scott. (1986) Effects of Instruction Based on the Madeline Hunter Model on Students' Achievement: Findings From a Follow Through Project. (In Press). Elementary School Journal.

Wolfe, P. (1982) Instructional Skills Observation Instrument. Napa County Office of Education, Napa, California.